UCLA

Network on Chip (NoC) : An Introduction

Outline

- What is a NoC
  - Motivation
  - Basics
  - Data Abstraction
- What it looks like
  - Hardware Components
  - Software Tools
- Interesting problems
- References

Introduction

- NoC : Is an on chip packet based communication system between blocks connected via routers
- Today application-specific systems on-chip (SoC) make extensive use of busses as the interconnect infrastructure
- However, in recent years research has shown that Network on-Chip (NoC) is likely to replace buses in future SoCs [1-4]

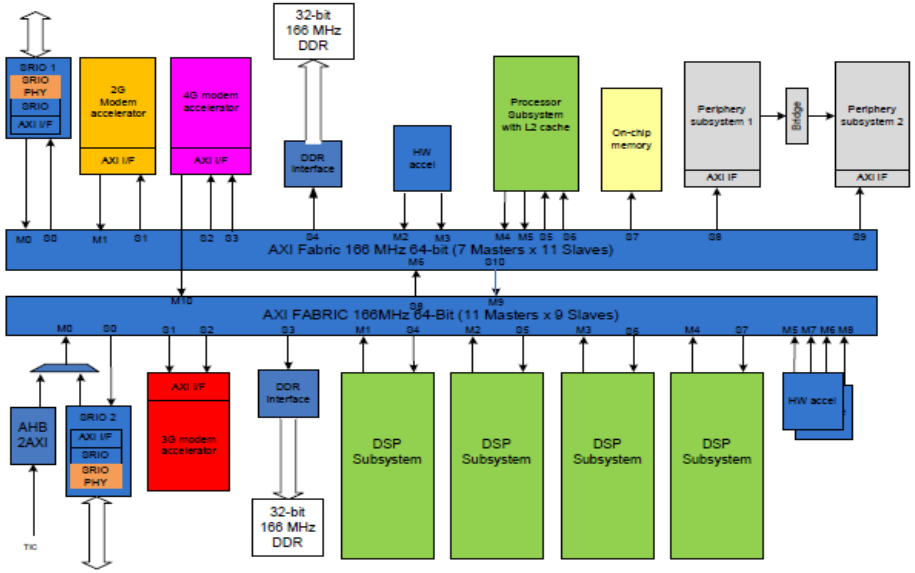


Fig. 1 : Bus based System Architecture

Motivation

- NoCs offer superior performance, power and area tradeoffs as the number of modules increases [1-4].
- Bus-based architectures have trouble scaling with increasing number of IP blocks and decreasing geometries
- NoC is a far easier-to-scale interconnect, with shorter, unidirectional, point-to-point wires [5]
- For QC ASICs NoC reduces interconnect by 30 – 40%

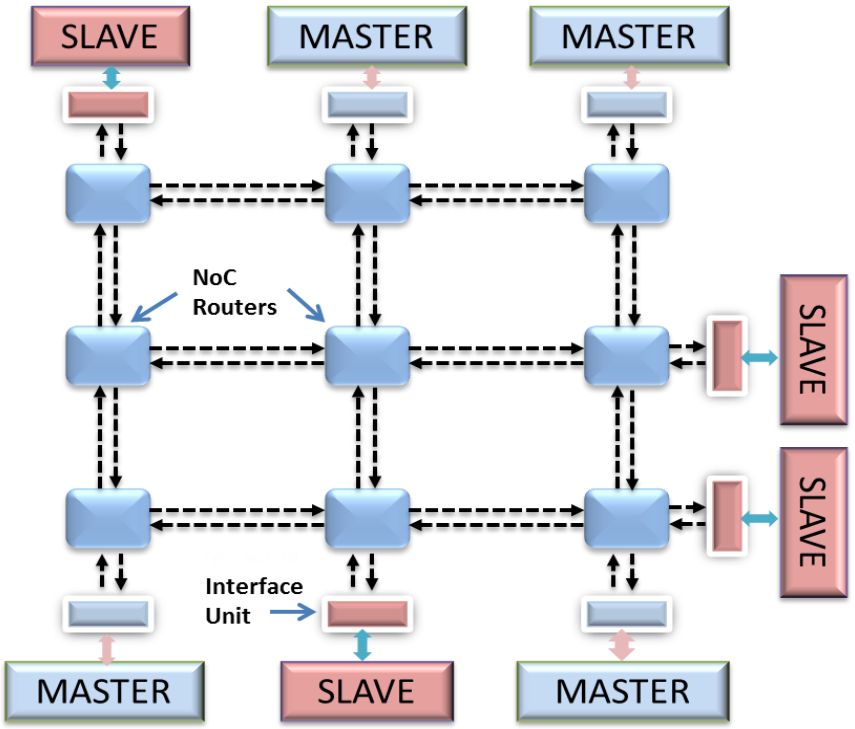
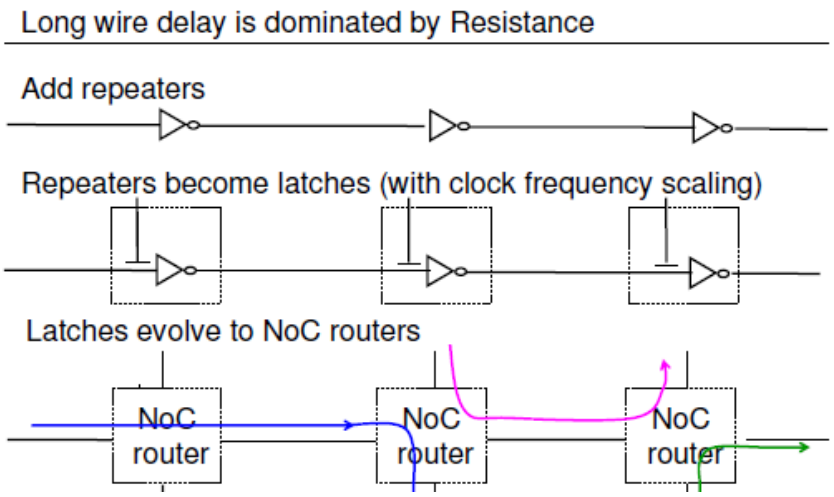
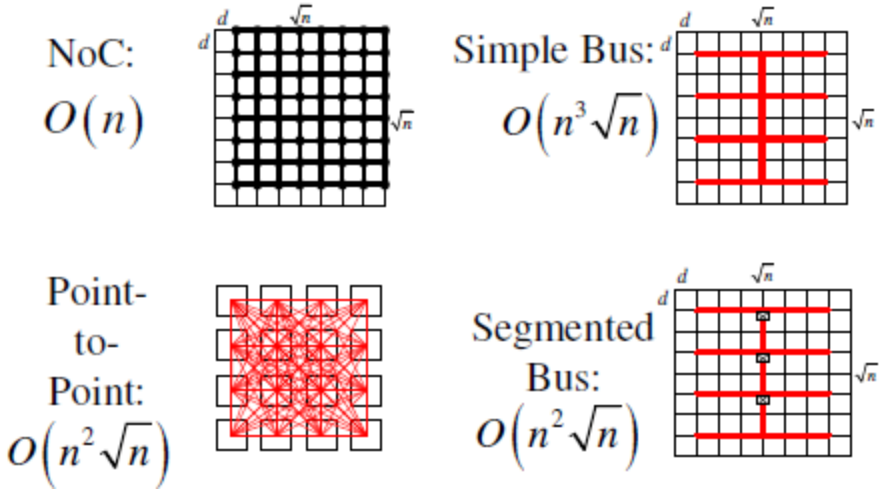


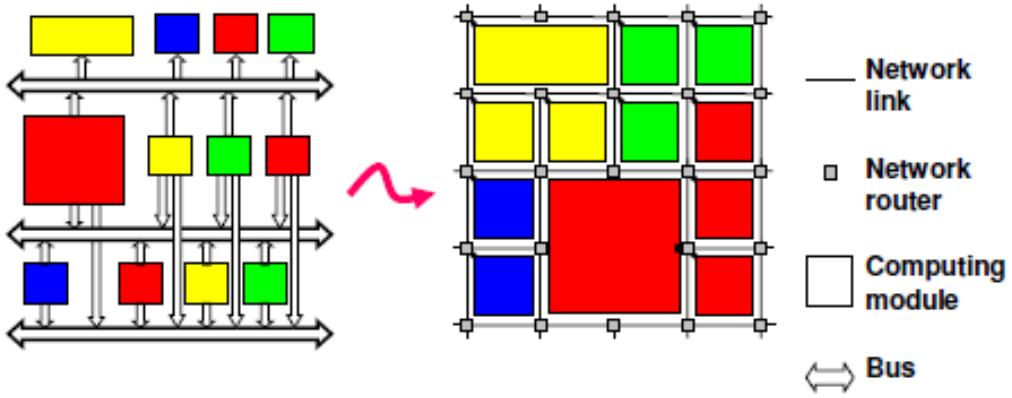
Fig. 2 : Generic 2-D NoC setup

NoC and Global Wire Delay





- Figure shows wire area cost for same performance
- Figures from [5] and [6]



Data Abstraction

- Data is transmitted or “routed” in the form of packets
- Just like regular networks, packets contain destination/source information etc.
- Functional simulation based trace data is often used for modeling during NoC generation flow

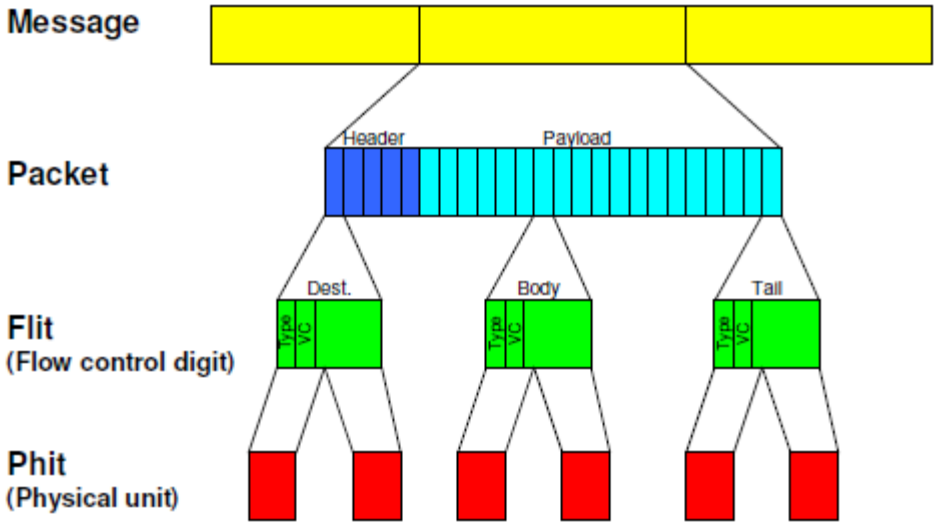


Fig. 3 : Levels of Data Abstraction

NoC Hardware

- Common building blocks :
  - Routers
  - Resizers
  - Clock-Synchronizer
  - NIU
  - FIFOs
  - Arbiters
  - Binners

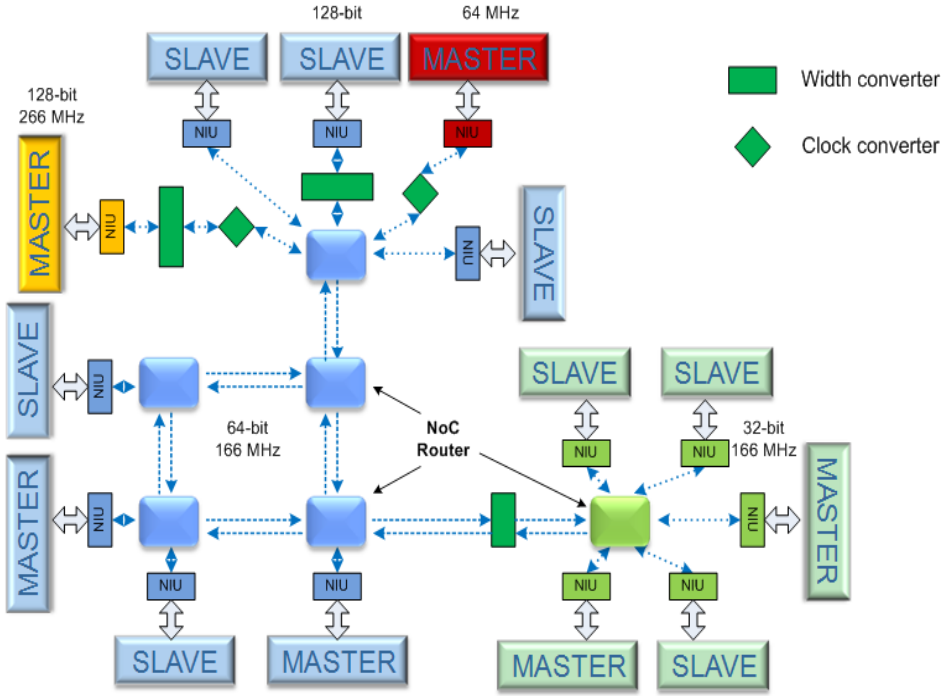


Fig. 4 : Closer look at a NoC

NoC Hardware : Routers

- Routers are the core of the NoC interconnect, typically lie on critical paths and area a major contributor to area
- A basic router is a MxN switch, with M-inputs and N-outputs
- Each input port is actually multiple signals (data, flow control, valid signals etc)

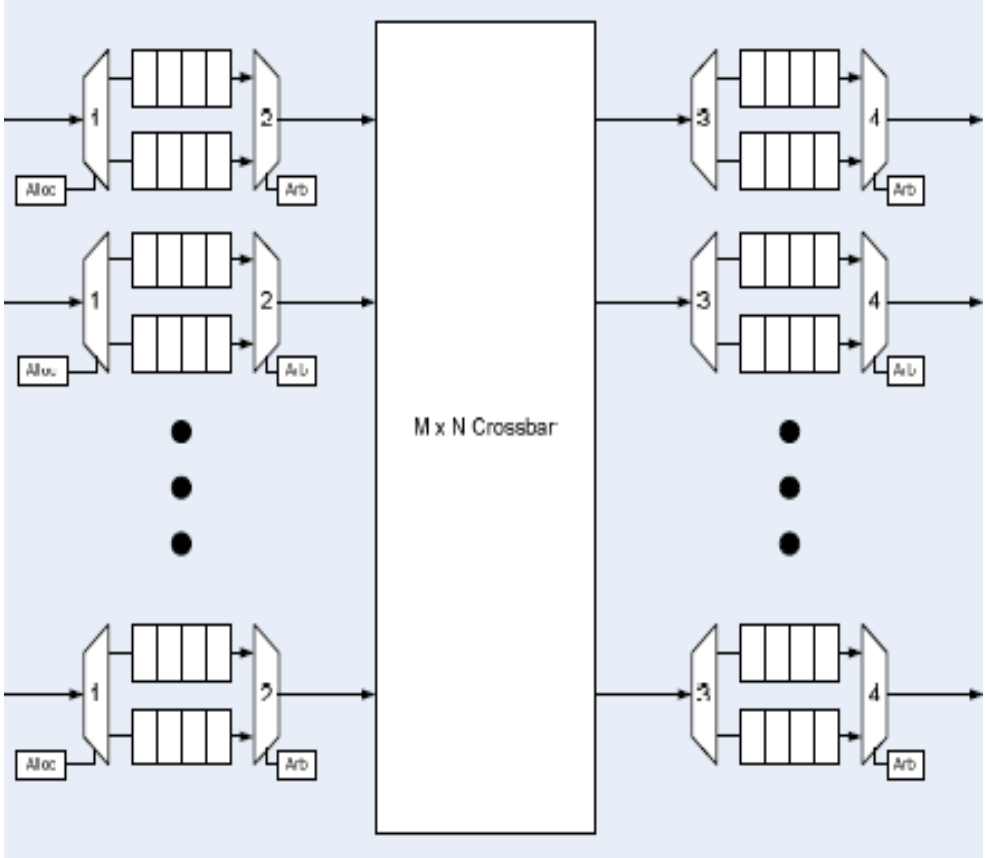
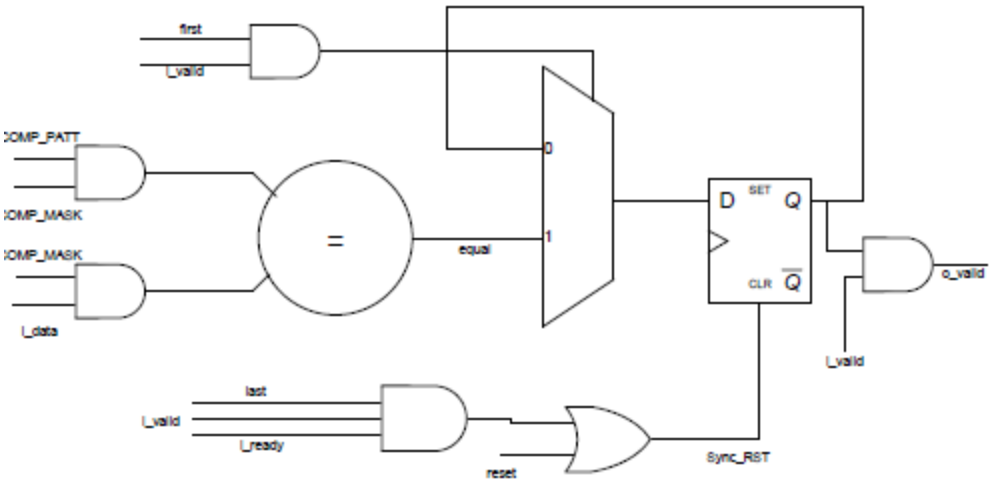


Fig. 5 : A basic NoC Router

NoC Hardware : “Binner”

- “Bins” packets into appropriate FIFOs/Queues
- “Binning” can be done on the basis of some subset of header information
- Allows for some notion of priority



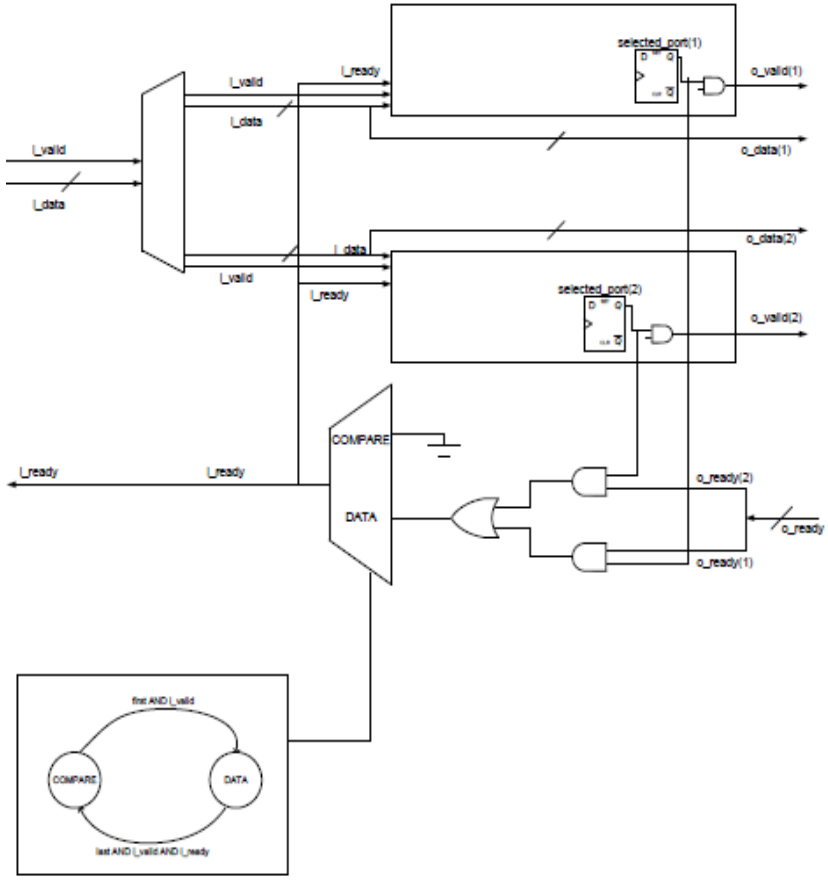


Fig. 6 : A basic NoC Binner Block

NoC Software Tools

- Topology Generation : often called NoC Placement, optimize multiple metrics, latency, bandwidth, size of routers
- System Simulation : to figure out traffic patterns, delay, latency, power etc. Also needed to figure out if constraints are met
- Route Optimization : figuring out optimal routing directions for packets from sender to destination
- Allowing for deadlock free network, is a challenge

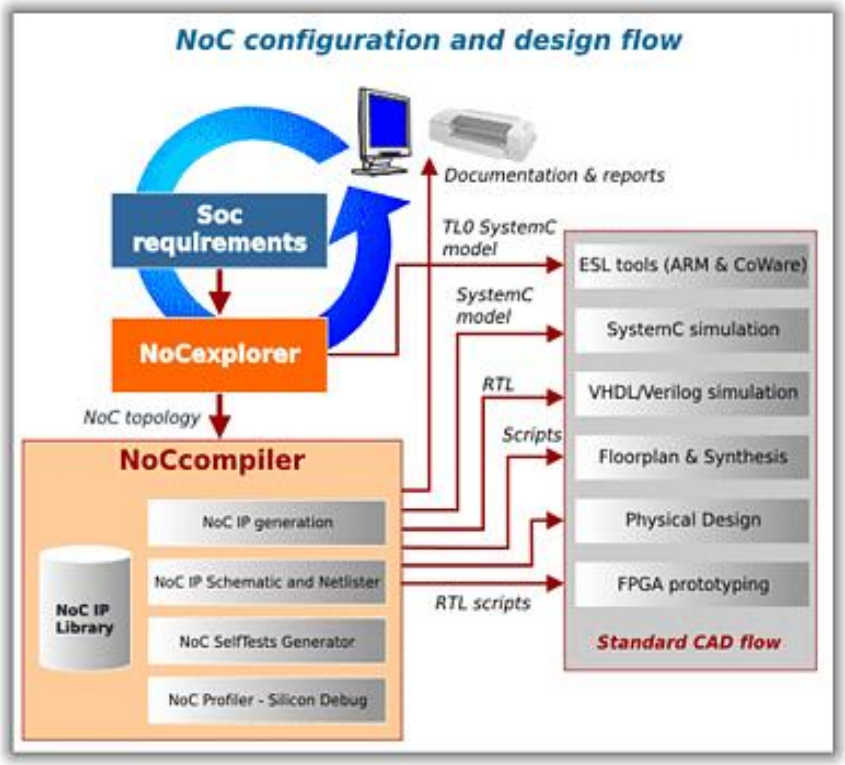


Fig. 7 : A basic NoC “Flow” from Arteris [7]

Interesting Problems : Placement

- Inputs : Traffic information, Module sizes, Link and Router Latency information (optional), Physical Constraints (optional)
- Come up with NoC router configuration and related core connectivity information
- Optimize on the basis of :
  - Bandwidth : Cores that “talk a lot” sit close to each other
  - Size of routers, router logic grows as O(n^2)
- Annealing based approach [8]
- How do we include floor-plan or physical information?

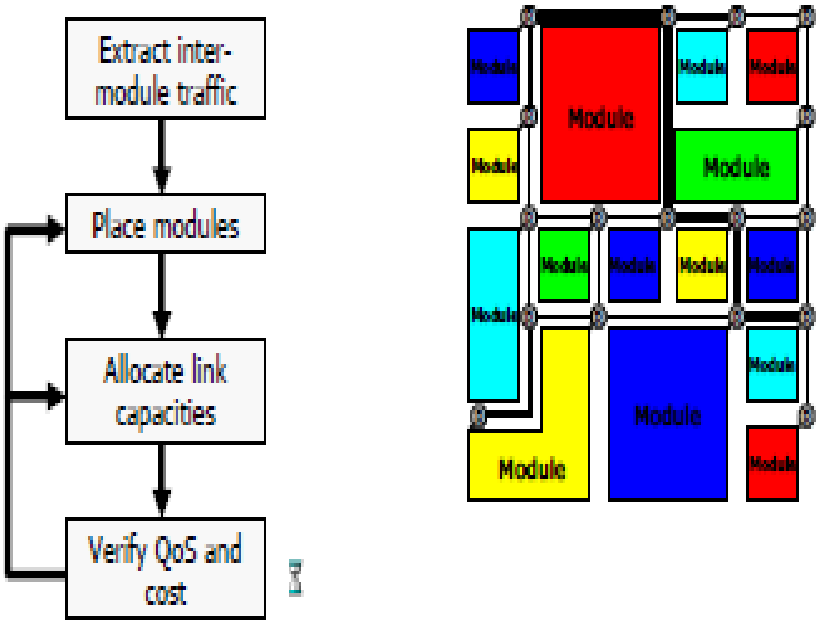


Fig. 8 : A Typical NoC Placement flow

Router Frequency and Width Assignment

- Assign frequencies and widths to routers in the NoC network, with the objective of minimizing the cost incurred due to :
  - Clock-Domain-Crossings
  - Width Conversion
- Finding globally optimal assignment has exponential growth in runtime – (#clock-domains)^(#routers)
- Constraint : Size and Frequency should satisfy bandwidth

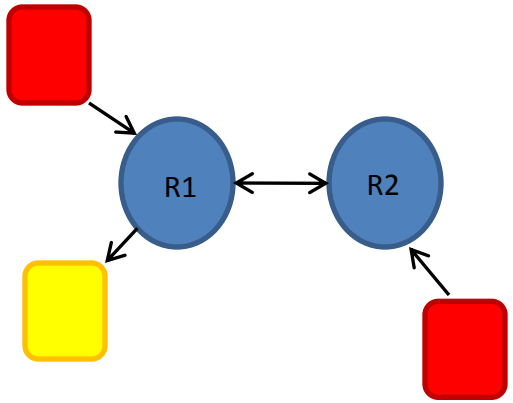


Fig. 9 : Toy router frequency assignment problem

VLSI CAD Problems Reframed for NoC[9]

- Floor-planning : NoC Placement
- Wire Routing : Message Routing
- Buffer/Inverter Sizing : Buffer (FIFO) Sizing
- Timing Closure : Link Bandwidth Capacity Allocation
- Simulation : Networking Simulation and Traffic Modeling

References

- [1] P. Guerrier and A. Greiner, "A Generic Architecture for On-Chip Packet-Switched Interconnections", Proc. Design, Automation and Test in Europe (DATE) 2000, 250-256
- [2] K. Goossens, J. Dielissen, and A. Radulescu, "A Ethereal Network on Chip: Concepts, Architectures, and Implementations", IEEE Design and Test of Computers, 2005, 414-421
- [3] E. Bolotin, I. Cidon, R. Ginosar, and A. Kolodny, "QNoC: QoS Architecture and Design Process for Network on Chip", Journal of Systems Architecture, Vol. 50, February 2004, 105-128
- [4] D. Bertozzi and L. Benini, "Xpipes: A Networkon- Chip Architecture for Gigascale Systems-on- Chip", Circuits and Systems Magazine, IEEE Volume 4, Issue 2, 2004, 18-31
- [5] E. Bolotin, I. Cidon, R. Ginosar, and A. Kolodny, "Cost Considerations in Network on Chip", Integration - the VLSI Journal, Volume 38, pp.19-42, 2004
- [6] Dally, W. J. and Towles, B. 2001. Route packets, not wires: on-chip inteconnection networks. In *Proceedings of the 38th Annual Design Automation Conference* (Las Vegas, Nevada, United States). DAC '01. ACM, New York, NY, 684-689. DOI= http://doi.acm.org/10.1145/378239.379048
- [7] Arteris : Network on Chip Company. http://www.arteris.com/
- [8] Beraha, R., Walter, I., Cidon, I., and Kolodny, A. 2009. The design of a latency constrained, power optimized NoC for a 4G SoC. In *Proceedings of the 2009 3rd ACM/IEEE international Symposium on Networks-on-Chip* (May 10 - 13, 2009). NOCS. IEEE Computer Society, Washington, DC, 86. DOI= http://dx.doi.org/10.1109/NOCS.2009.5071449
- [9] Technion NoC Tutorial : http://webee.technion.ac.il/matrics/talks/Technion%20NOC%20Tutorial%20Part%20Four.pdf